DOCUMENT RESUME

ED 297 031 TM 012 010

AUTHOR Straus, Murray A.
TITLE Validity of Aggregating Individual-Level Survey Data for Macrosociological Research Using U.S. States.
INSTITUTION New Hampshire Univ., Durham. Family Research Lab.
SPONS AGENCY Northwest Regional Educational Lab., Portland, Oreg.
PUB DATE Oct 87
GRANT T32-MH15161
NOTE 21p.; For a related document, see TM 012 000. Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Concurrent Validity; Construct Validity; Demography; Meta Analysis; *National Surveys; *Social Science Research; *State Norms; State Surveys; *Statistical Analysis; Violence
IDENTIFIERS *Aggregation (Data); General Social Survey; National Family Violence Survey; *States (Geopolitical Regions)

ABSTRACT

The feasibility of employing survey data that uses the states of the United States as the societal units in macro-level sociological research was studied. The range of issues available for research opened to investigation would be greatly expanded if it were possible to aggregate national survey data to produce state-level statistics for variables such as approval of violence and gender role attitudes. If such data were to be available, they could be used to investigate issues such as whether state-to-state differences in the degree of approval of violence explain part of the huge differences among states in the rate of violent crimes. The validity analyses reported in this paper were conducted using data from the 1975 National Family Violence Survey, the 1985 National Family Violence Resurvey, and the 1972-84 cumulative General Social Survey. Analyses provide information on concurrent validity, construct validity, validity of specific variables, and multi-indicator indexes. It is concluded that aggregate survey data should be avoided unless there are strong reasons to use such data despite the problems. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED297031

Validity of Aggregating Individual-Level Survey Data for Macrosociological Research Using U.S. States*

Murray A. Straus
Family Research Laboratory
University of New Hampshire, Durham, NH 03824

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
MURRAY A. STRAUS

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 012 010

VALIDITY OF AGGREGATING INDIVIDUAL-LEVEL SURVEY DATA FOR MACROSOCIOLOGICAL RESEARCH USING U.S. STATES*

Murray A. Straus
Family Research Laboratory, University of New Hampshire
Durham, NH 03824 (603) 862-2594

OBSTACLES TO USE OF NATIONAL SURVEY DATA 2
Inadequate Size Sample
Limited Number of States
Sample Design Not Usually Appropriate
METHOD 3
Data
Concurrent and Construct Validity Analyses
Ambiguity In Criteria For Validity
THE 1975 NATIONAL FAMILY VIOLENCE SURVEY 5
Sample
Concurrent Validity
Construct Validity
THE 1985 NATIONAL FAMILY VIOLENCE RESURVEY 6
Sample
Effect Of State N On Concurrent Validity
Construct Validity
Effect of Sample Size
THE GENERAL SOCIAL SURVEY 8
Concurrent Validity
Construct Validity
Sexual Liberalism Index
Violence Approval Index
SUMMARY AND CONCLUSIONS 10
Concurrent and Construct Validity
Conclusions

---

* Paper to be submitted for presentation at the 1988 annual meeting of the American Sociological Association. This research is part of the State and Regional Indicators Archive and research program of the Family Research Laboratory, University of New Hampshire, Durham, NH 03824. A program descriptions and publications list will be sent on request.

It is a pleasure to acknowledge the support of these organizations and to express appreciation to Kersti Yllo who created the state-level data file for the 1975 National Family Violence Survey, to Christine Smith for creation of the state-level file for the 1985 survey, to James A. Davis and Tom W. Smith who, despite misgivings, authorized the creation of state-level statistics from the General Social Survey, and to David Jaffee for converting those statistics into a file of state-level variables. Funding for this research was provided by the University of New Hampshire and National Institute of Mental Health (grant T32 MH15161).

The objective of this paper is explore the feasibility of employing survey data in macro-level research which uses the states of the United States as the societal units. A previous paper (Straus, 1985a) presented theoretical and empirical evidence on the validity of using the states of the United States for macro-sociological research. However, such research faces a practical limitation if the data are restricted to census and other government statistics. The range of issues which could be investigated would be greatly expanded if it were possible to aggregate national survey data to produce state-level statistics for variables such as approval of violence and gender role attitudes. If such data were to be available and valid, it could be used to investigate issues such as whether state-to-state differences in degree of approval of violence explains part of the huge differences between states in the rate of violent crime.

OBSTACLES TO USE OF NATIONAL SURVEY DATA

Although the possibility of employing the vast storehouse of national survey data which has accumulated since World War II is attractive, the reality is much more limited.

Inadequate Size Sample

Only a small proportion of the huge number of surveys available in data archives such as the Interuniversity Consortium For Political and Social Research at the University of Michigan, the Roper Center at the University of Connecticut, and the Louis Harris Political Data Center at the University of North Carolina are usable for state-by-state research because the typical national survey includes about 1,200 respondents. Even if there were an equal number of cases drawn from each state (which is never the case), there would be only 24 cases per state. For this reason, the surveys used for the research described in this article were selected because each has a sample size that is at least double the size of the typical national survey. Moreover, even in large N surveys there will be states represented by as few as two or three respondents.

Adding to the problem of sample size is the fact that there is no clear criterion for determining the minimum number of cases. It depends on the use being made of the data. If, for example, the data is being used to compute correlations between median education (as determined by the census) and the percent in each state holding a certain attitude, much lower n's can be tolerated than if the purpose is to present state-by-state descriptive statistics. While there may be exceptions, such as presenting such data for the ten largest states, state-level statistics produced from individual-level surveys should be used only to investigate relationships between variables by techniques such as correlation and regression, and should <u>not</u> be used to present descriptive statistics on the percent of people in a given state or states who hold a certain attitude, or who have certain characteristics.

Publications reporting results based on data aggregated from surveys should focus on measures of association (cross-tabs, ANOVA, correlation, regression, etc.). It is generally best to avoid reporting the score for a

particular state because (for the reasons given above) there is a large risk of error connected with any one data point. --

Limited Number of States

One method of overcoming the problem of states represented by few respondents is to eliminate the states with a low N. However, unless the survey is extremely large, this results in a drastic reduction in the number of states available for analysis. For example, a survey with an N of 5,000 might include respondents from 40 states. But even if one sets a minimum number of cases as low as 50 per state, the sample size is likely to drop to 25 states.

Sample Design Not Usually Appropriate

Even when the survey is large enough to aggregate to the state-level, another problem is almost always present. Most national surveys are designed to be representative of each region, but not necessarily of the states within each region. For example, as the first step in the sample design, all the counties in the North East might be randomly sampled. However, in any one state, only one or two counties might be selected. Both of those could be rural counties, or both highly urban counties. Thus, it is possible to have a "sample" of Massachusetts in which the Boston metropolitan area is not included at all.

The obstacles listed above are formidable. On the other hand, there may be important issues for which no other data is available. This was the situation that led Yllo and Straus (Yllo, 1983a, Yllo and Straus, 1984) to explore the possibility of creating state-level variables from an individual-level survey. The results (to be described below) were sufficiently promising to encourage the more extensive methodological analysis presented in this paper.

METHOD

Data

The validity analyses reported in this paper were carried out using data from three national surveys: the 1975 National Family Violence Survey (Straus and Gelles, 1980; Straus, Gelles and Steinmetz, 1980), the 1985 National Family Violence Resurvey (Straus and Gelles, 1986), and the 1972-84 cumulative General Social Survey (Davis and Smith 1985). Each of these surveys use relatively large size samples, and each have been the basis for numerous publications. Each survey will be described in more detail in the sections where the findings from that survey are presented.

The state-level variables were created by computing the percentage of respondents in each state who expressed a certain opinion or who reported a certain behavior or socioecomic characteristic. An example of using individual attitude data to create a measure of social norms for each state, is the percentage of respondents in Alabama, Alaska, Arkansas, etc who endorse the death penalty. Examples of using individual behavior measures to create behavioral structure measures for each state include

the percent of the population each state who own a handgun, who slapped their spouse during the previous 12 months, or who drank more than a certain number of ounces of liquor during the week of the survey.

Concurrent and Construct Validity Analyses

Both "concurrent validity" and "construct validity" (Cronbach, 1970; Nunnally, 1978; Straus, 1964) analyses of these state-level variables were conducted. The concurrent validity analyses consisted of determining the degree to which the state-level estimates based on the survey data are correlated with the same variable as given in the U.S. census; for example, the correlation of the median income reported by respondents in each state with the median income as reported in the census. The construct validity analyses investigated the degree to which state-level variables created from individual-level surveys produce findings which are consistent with theoretical or empirical propositions. For example, if the percent of respondents in each state who report use of alcohol is correlated with the death rate from cirrhosis of the liver, that provides evidence of construct validity because it is known that heavy drinking is the major cause of cirrhosis.

Ambiguity In Criteria For Validity

Concurrent Validity. Remarkable as it may seem, there are no established standards for judging concurrent validity coefficients. Inspection of several psychometrics texts revealed that almost none give numerical figures, nor does the Standards For Educational and Psychological Tests and Manuals published by the American Psychological Association. Perhaps the reason is that the assessment of validity is a complex issue that is best approached multidimensionally (see for example, Brindberg and Kidder, 1982; Campbell and Fiske, 1959). Nevertheless, some numerical frame of reference can be helpful. Cronbach (1970) is one of the few authors who provides this. His Table 5.3 "Illustrative Validity Coefficients" includes 18 coefficients for widely used tests and sub-tests. My tabulation of these coefficients shows that they range from .08 to .77, with a mean of .37. Cronbach comments "It is unusual for a validity coefficient to rise above 0.60...."

The absence of established standards for judging concurrent validity is even more of a problem sociology. In fact, sociological research reports rarely present any validity evidence at all (Straus, 1964). Sociologists place great importance on the representativeness of the sample, and seem to implicitly assume that if the sample is representative, the measures used in studying that sample are valid.*1

In view of the absence of alternative empirically or theoretically derived criteria, it was decided to compare the validity coefficients computed for this paper with the mean of coefficient of .37 derived from Cronbach's Table 5.3.

Construct Validity. Construct validity refers to the extent to which the pattern of association between the measure in question and other variables follows a pattern that is consistent with theoretical or empirical knowledge (Cronbach, 1970; Nunnally, 1978; Straus, 1964). Thus,

a measure of the caloric intake should be correlated with feeling hungry, based on the theory that the subjective experience of hunger is caused by lack of food intake. Of course, the correlation will be less than 1.00 because there are other factors which also influence subjective feelings of hunger. There is even more ambiguity as to the size of the coefficient which will be taken as evidence of construct validity than there is for concurrent validity. This is inherent in the process. If the theory being tested with the new measure specifies a close linkage between the independent and dependent variable, then a large correlation is needed; but if (as in most theories) only a weak bivariate relation is posited because of the numerous other factors which are involved, then low correlations, provided they are statistically significant, support the construct validity of the measures used to test the theory.

THE 1975 NATIONAL FAMILY VIOLENCE SURVEY

The first exploration of the possibility of using individual-level survey data for state-level macro-sociological research grew out of the importance of testing the theory that "wife-beating" is one of many socially patterned mechanisms which serve to keep women subservient to men (Yllo 1983a,b Yllo and Straus (1984). Since wife-beating rates were not available for societal units such as nations, cities, or states, it was decided to create estimated rates for each of the states included in the National Family Violence Survey (NFVS).

Sample

The NFVS is an interview study of a nationally representative sample of 2,143 adults who were married or living with a partner of the opposite sex in 1975. The survey included respondents in 36 states. The number of cases per state ranges from 3 to 167, with 8 states represented by less than 20 cases per state. The data tape is available from the Interuniversity Consortium for Political and Social Research (Straus and Gelles, 1980, ICPSR study 7733).

Concurrent Validity

Concurrent validity was investigated by computing the correlation between five state-level variables created by aggregating the survey data with five census variables which measure approximately the same characteristic. These correlations were computed for the entire set of 36 states, and then replicated after deleting the six states represented by less than 20 respondents.

(Table 1 about here)

The correlations between the survey data estimates and the census data using all 36 states ranged from .13 to .68, with a mean of .46 (Table 1). The correlations using the 30 states with N's of at least 20 cases were much higher --- from .24 to .77, with a mean of 58. Thus, even when the state-level variables are include states with fewer than 20 respondents, the correlations exceed the average reported in psychometric validity studies, as summarized in the preceding section of this paper.

The validity coefficients for the 30 states with 20 or more cases are also remarkably high in view of the fact that even the largest states are represented by only about 100 cases, 12 of which were states in the 20 to 50 respondent range. In addition, the census variables use a different base population than the survey data, i.e. all adult males (census) versus married males living with spouse (survey). Consequently, even complete enumeration data would not produce a perfect correlation.

Construct Validity

Some evidence of "construct validity" is provided by Yllo and Straus's analyses (Yllo, 1983a,b, Yllo, 1984) of the relationship between a "Status of Women" index computed for each state and the rate of wife-beating computed for each state from the data in this survey. The results are consistent with hypotheses based on conflict theory, and therefore provide at least some evidence of the "construct validity" of the incidence rates based on aggregating survey data.

THE 1985 NATIONAL FAMILY VIOLENCE RESURVEY

Sample

Interviewing for the National Family Violence Resurvey (NFVR) was conducted in the summer of 1985. A total 6,002 persons age 18 and over were interviewed, and a number of reports have been published or are in press (Straus and Gelles, 1986, Kaufman Kantor and Straus, 1987; Gelles and Straus, forthcoming, 1988). The large sample size and the fact that the survey included respondents in all 50 states and the District of Columbia made it possible to conduct more extensive analysis than was could be done with the original family violence survey.*2 For example, the number of cases per state ranges from 7 to 570. This range permits a more systematic investigation of the effect of the number of respondents per state. An important difference between this survey and the other two surveys is that the sample was drawn by random methods within each state. Consequently, it is hypothesized that the validity of state-level variables computed from this survey is greater than the validity of the state-level variables computed from the other two survey, neither of which were designed to be representative of specific states.

Effect Of State N On Concurrent Validity

(Table 2 about here)

The coefficients in the three columns of Table 2 headed "Minimum Number of Cases Per State" provide data on the extent to which validity is increased by requiring a certain minimum number of cases per state. Specifically, these coefficients permit a test of the hypothesis that the validity of state-level variables will increase in proportion to the number of respondents per state.

Contrary to this hypothesis, the correlations in the first column (which uses only states represented by at least 100 respondents) are not much greater than those in the second column (which uses states with as

few as 25 respondents), or even the third column (which includes all 50 states and the District of Columbia, regardless of the number of respondents in each state. The reasons for these unexpected results are not clear. Perhaps there is a trade-off between grains from a more adequate sample size and losses from a more restricted number of states.

Comparison of the average validity coefficients in the bottom row of Table 2 with the average in the bottom row of Table 1 shows that the 1985 coefficients are over 50% higher than the coefficients based on the 1975 survey. The greater validity of the state-level estimates from the 1985 study is probably a result of the combined effect of the much larger sample studied in 1985 and the fact that this sample was selected to be representative of each state in the study.

Construct Validity

(Table 3 about here)

Table 3 relates 12 state-level variables created by aggregating the National Family Violence Resurvey to the state-level to variables based on public data sources such as the US Census and the Vital Statistics of the United States. In contrast to the correlations in Table 2, the dependent variables in Table 3 are not intended to be measure of the same variable as was measured in the NFVR. Rather the pairs of variables correlated in Table 3 were chosen on the basis of theoretical assumptions. Consequently, these correlations provide a means of examining the "construct validity" of the NFVR variables.

Effect of Sample Size. The first of the two columns of validity coefficients at the right of Table 3 makes use of all 50 states and the District of Columbia, whereas the second of the two columns uses only the 36 states represented by 100 or more respondents. Comparison of the two columns of correlations in Table 3 shows that, without exception, the correlations in the second column are higher than those in the first column. Since the second column was computed using the 36 states represented by 100 or more respondents, this would be an unremarkable finding if it were it not for the fact that it is inconsistent with the results in Table 2. The Table 2 results show only small differences according to which set of states is used. Since no explanation has been developed for the Table 2 findings, and since those based on Table 2 seem to be more plausible, it seems safest to conclude that the most valid way to analyze these data may be to restrict the analysis to the 36 states which are represented by at least 100 respondents.

Homicide. Part A of Table 3 shows the relationship of three state-level variables based on the NFVR to the state homicide rate. The findings are consistent with what is known about the causes of homicide. Specifically, the first row shows that the larger proportion of the population of a state who regard it as permissible for a husband to hit his wife, the higher the homicide rate. The second row shows that the higher the average score on a test intended to measure overt aggressive acts, the higher the homicide rate. The third row shows that the higher the percent of the population who are black, the higher the homicide rate. The latter correlation is particularly strong, which is consistent with the fact that the homicide rate among the black population is several

times greater than among the white population (Curtis, 1975, Plass and Straus, 1987).

Alcoholism. Part B of Table 3 follows the same logic for state-level variables based on self-reported alcohol use. All eight of the self-report based measures of drinking were found to be correlated with estimates of alcoholism based on the death rate for cirrhosis of the liver.

Poverty. Part C of Table 2 shows evidence of the validly of state estimates based on the income question in the NFV Resurvey. The correlations are lower than the concurrent validity correlations for income in Table 2 because the dependent variable is the percentage of children living in poverty, not the mean or median family income.

Stress and Psychological Problems. Part D of Table 2 examines the validity of three measures of psychological well-being of the population in each state. The question is whether one can use the results of this survey to measure state-to-state differences in the average level of psychological well-being. The first row of this Part D shows that subjectively experienced stress is not significantly associated with the frequency with which "stressful events" occur in each state. However, the next two rows of Part D show that, as predicted, the higher the rate of stressful "life events" in a state, the higher the rate of self-reported depressive symptoms (such as feeling sad, suicidal thoughts), and the higher the rate of psychosomatic complaints (such as headaches, cold sweats).

THE GENERAL SOCIAL SURVEY

This section presents the results of validity analyses using the cumulative data file for the General Social Survey (GSS). The extremely large sample size (over 15,000 cases) and the wide range of topics covered in the GSS made it possible to compute numerous validity coefficients. The GSS surveys used for these analyses were conducted in 1972 through 1978,1980, and 1982 through 1984. A total of 15,579 cases are in the individual-level files.*3

Despite the large number of cases, the respondents were drawn from only 41 states and the N per state ranges from lows of 16 for Mississippi and 26 for Rhode Island to highs of 1,005 for New York and 1,583 for California. There are 34 states represented by 50 or more respondents.

Concurrent Validity

(Table 4 about here)

Thirty eight of the individual-level variables from the GSS were used to create state-level variables which correspond, at least partly, to census variables. Correlations between these 38 GSS variables and 60 census variables are presented in Table 4.

Effect of Sample Size. The first column Table 4 gives the concurrent validity coefficients when all 41 states in the GSS are used, including

seven states represented by less than 50 respondents. The middle column of correlation coefficients excludes these seven states, and the column at the far right uses only the 23 states that were represented by 200 or more respondents. The last row of Table 4 was computed to test the hypothesis that the larger the number of cases used to create a state-level variable, the more valid the variable. It gives the average validity coefficient for these three "samples." Surprisingly, there is almost no gain in the average validity coefficient when the seven states represented by less than 50 respondents are excluded (r = .34 and .36). Moreover, there is only a moderate gain in validity when states with less than 200 respondents are excluded (r = .45). This replicates the findings from the same type of analysis done using the National Family Violence Resurvey data.

Validity of Specific Variables. Another puzzle is why the validity coefficients vary so greatly from variable to variable. For example, the rows in Part A of Table 4 for variable GS1T1 show a high level of concurrent validity for the GSS based measure of Civilian Labor Force (CLF) participation rate, and the rows in Part C for variable GS3T3 show an even higher validity for the GSS based measure of divorce. On the other hand, the coefficients in Part D of Table 4 are low, some of them near zero.

Comparison with Family Violence Survey. An important finding of this study is revealed by comparing the average validity coefficients for the GSS variables (last row of Table 4) with the average validity coefficient for the 1985 National Family Violence Survey (last row of Table 2). Consistent with the hypothesis posed earlier, the valid of state-level variables derived from the GSS is substantially lower, even though the GSS sample is many times larger than the 1985 family survey sample. It seems plausible to attribute the lower validity of the GSS state-level variables to the fact the sample for the GSS is not intended to be representative of each state in the sample. The other side of the coin, however, is also important; namely that despite this, the concurrent validity coefficients are as high as they are. Moreover, as will be shown in the next section, the construct validity of variables computed from the GSS is as higher or higher than validity coefficients found anywhere in the sociological or psychometric literature.

Construct Validity Of Multi-Indicator Indexes

The original reason for creating state-level variables from the GSS was to measures of constructs needed to test certain theories concerning the social causes of rape and other violence: a measure of normative support for violence, and a measure of sexual liberalism or tolerance. The wording of the questions and response categories for the items in each index and the method of combining indicators to form the composite index are given in State and Regional Indicators Archive Codebook (Straus, 1987).

Sexual Liberalism Index. The indicators making up this index are 20 attitude items, each of which is in the form of the percent of respondents in a state who express a favorable attitude toward abortion, allowing homosexuals to teach in a college, permitting teen agers to have access to

contraceptives, and favoring sex education in the schools. The index has an alpha coefficient of reliability of .96. The index was used in Jaffee and Straus (1986) and is described in footnote 3 and in the methodology section of that paper.

Jaffee and Straus (1986) found a correlation of .50 between this index and the circulation rate of newsstand sex magazines in each state. Since these two variables have a very different origin, and since one refers to attitudes and the other to overt behavior (purchase of sexually explicit materials), this correlation is evidence suggesting both construct and discriminant validity.

Violence Approval Index. Responses to 14 GSS questions were used to compute this index. Each item consists of the percentage of respondents in each state who are opposed to requiring gun permits, favor the death penalty, believe that more should be spent on the military, and approve of hitting another person under a variety of different conditions. The Violence Approval Index has an alpha coefficient of reliability of .67.

Baron, Straus and Jaffee (1987) found a correlation of .40 between this index and an index designed to measure "legitimate violence" (Straus, 1985b; Baron and Straus, n.d.). The two indexes are intended to measure the same underlying concept -- socially permissible and/or approved violence. However, they use entirely different data: attitudes in the case of the Violence Approval Index described above and laws and behavior in the case of the "Legitimate Violence Index." For example, the Violence Approval Index uses the percent of respondents in each state who oppose requiring gun permits, and the Legitimate Violence Index uses a parallel behavioral measure -- the circulation rate per 100,000 population of gun magazines. Since the Legitimate Violence Index is based on "objective" data, as contrast to the Violence Approval Index which is based on GSS attitude data, the correlation suggests the validity of both measures.

Baron, Straus, and Jaffee's test of the "cultural spillover" theory provides evidence of construct validity for the Violence Approval Index. They hypothesized that the higher the support for culturally permissible violence, the higher the rate of criminal violence. This hypothesis was supported using both the Violence Approval Index computed on the basis of the GSS and the Legitimate Violence Index computed from public record data such as state laws and state expenditures per capita on the National Guard.

SUMMARY AND CONCLUSIONS

A previous paper provided evidence of the validity of using the States of the United States as units of society for macro-sociological research (Straus, 1985a). However, a practical limitation arises despite the large amount of census and other public record data on the states because many important issues can only be investigated if it is possible to aggregate individual-level survey data as a means of measuring the properties of states.

Macro-level variables are frequently created from national surveys for use in cross-national comparative research but have not been used for cross-state comparative research. Some of the obstacles and problems

connected with using survey data to measure properties of states are discussed, including an inadequate number of cases per state in a typical national survey, the fact that national surveys using include respondents from only about 30 states, and the fact that the sample for most national surveys is not designed to be representative of each state where respondents are interviewed.

The main part of the paper reports empirical studies of the validity of state-level variables created by aggregating individual-level surveys. The empirical analyses use three large and well proven surveys to create state-level variables. The procedure was to compute the percent of respondents in each state who expressed a certain opinion, or who reported having certain characteristics. Two types of state-level variables were computed: variables needed for substantive analyses, such as the rate of wife-beating, and variables which correspond to census or other public data. The latter provided a means of estimating the "concurrent validity" of state-level variables created by aggregating individual responses for each state, and the former provide a means of investigating the "construct validity" of state-level variables created from individual-level survey data.

Concurrent and Construct Validity

Almost all the concurrent validity coefficients (correlations of state-level variables created by aggregating individual-level survey data with measures based on census and other public record data) exceed the average concurrent validity coefficient reported in the psychometric literature (.37), most by a considerable margin. For the survey best suited to creating state-level variables, the average validity coefficient was found to be .77.*[4]

The findings for construct validity provide considerable evidence supporting the idea that conceptually valid variables can be created for the states of the United States by aggregating individual-level survey data. The construct validity of variables based on the General Social Survey is particularly encouraging because this is such a widely used data set and because the variables measure so many key aspects of American society.

In general, the size of both the concurrent validity and the construct validity coefficients is remarkable because the samples for two of the three surveys used to create the state-level variables were not drawn by methods intended to create a valid sample within each of the states that happened to fall in the survey, and because for each of the three surveys a substantial number of states were represented by 25 or fewer respondents.

Conclusions

The discussion up to this point has tended to interpret the results as indicating a surprisingly high level of validity for state-level variables computed by aggregating individual-level survey data. However, while the average is higher than was expected, there is considerable variability. Moreover, even the average needs to be looked at from

another perspective. A concurrent validity coefficient of .50, for example, can be regarded as remarkable high or remarkably low. Should it be taken as evidence that the glass is half empty or half full? In my opinion, both these emphases are misleading. Rather, one needs to be aware of both the strengths and the limits of any set of measures. The results reported in this paper are a step in that direction.

What do these findings tell us about the practical issue of whether to proceed with using variables created from individual-level surveys. On the one hand, the evidence of concurrent and construct validity is encouraging. On the other hand, the typical national survey is not designed to produce valid state-level findings and has an inadequate number of cases per state. Given these conflicting consideration, there can be no general recommendation. Perhaps the most prudent approach is to conclude that aggregated survey data should be avoided unless there are strong reasons to use such data despite the problems. Among those reasons are lack of a feasible alternative, the uniqueness and importance of the issue being investigated, and the results of methodological analyses which provide information on the level of confidence that one can have in a particular set of data.

FOOTNOTES

1. The situation is almost the opposite in psychology. Relative to sociologists, psychologists pay much more attention to the validity of the measures and seem to implicitly assume that if the measure is valid, the sample is not crucial.

2. Respondents were selected by four methods. A national probability sample of approximately 4,000, oversamples to increase the number of black and hispanic families, and an oversample to increase the N's for certain states to 100 per state. The oversamples have been weighted to enable all 6,002 cases to be used as a nationally representative sample. The state-level variables used in this paper are based on this weighted total sample.

3. I would like to express my appreciation to James A. Davis[5] and Tom W. Smith for providing the state-by-state statistics used in this paper. These statistics had to be computed by the staff of the National Opinion Research Center because the GSS sample is not designed to be representative within each of the states. Consequently, the public use data tape does include a state identification code.

4. Many sociologists will be inclined to regard an average validity coefficient of .37, or even .77, as indicating the low standard prevailing in psychology and a low validity of state-level variables created by aggregating survey data. That may well be true, but it does not follow that sociologists do better, or that .77 is a low average validity coefficient. In fact, such conclusions are almost surely not warranted because there is no basis for comparison. Previous analyses reveal that sociologists almost never report validity coefficients for their measures (Blalock, 1979; Straus, 1964; Straus and Brown, 1978). That situation continues to this day. In the 1986 volume of the American Sociological Review, not one investigator reported a validity coefficient.

REFERENCES

Baron, Larry and Murray A. Straus. n.d. Rape In American Society: The Social Causes of Sexual Violence. Book in preparation.

Baron, Larry, Murray A. Straus, and David Jaffee. 1987. "Legitimate Violence, Violent Attitudes, and Rape: A Test of the Cultural Spillover Theory." Annals of the New York Academy of Science.

Blalock, H. M. 1979. "Measurement and Conceptualization Problems: The Major Obstacle to Integrating Theory and Research." American Sociological Review 44:881-894.

Brinberg, David and Louise H. Kidder, eds. 1982. New Directions for Methodology of Social and Behavioral Science: Forms of Validity in Research. San Francisco, CA: Jossey-Bass.

Campbell, Donald T. and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." Psychological Bulletin 56:81-105.

Cronbach, Lee J. 1970. Essentials of Psychological Testing. Third Edition. NY: Harper & Row.

Curtis, Lynn A. 1975. Violence, Race, and Culture. Lexington, MA: Heath, Lexington Books.

Davis, James A., and Tom W. Smith. 1985 General Social Surveys, 1972-1984: Cumulative Codebook. Chicago, IL: National Opinion Research Center. Distributed by the Roper Public Opinion Research Center, University of Connecticut, Storrs, CT.

Gelles, Richard J. and Murray A. Straus. 1988 forthcoming. Intimate Violence. New York: Simon and Schuster.

Jaffee, David, and Murray A. Straus. 1986. "Sexual Climate and Reported Rape: A State-level Analysis." Archives of Sexual Behavior 16:107-123.

Kantor, Glenda Kaufman and Murray A. Straus. 1987. "The Drunken Bum Theory of Wife-Beating." Social Problems 34:213-230.

Nunnally, Jum C. 1978. Psychometric Theory. Second Edition. NY: McGraw-Hill.

Plass, Peggy S. and Murray A. Straus. 1987. "Intra-Family Homicide in the United States: Incidence, Trends, and Differences by Region, Race, and Gender." Paper presented at the Third National Family Violence Research Conference, University of New Hampshire, Durham, NH, July 6.

Straus, Murray A. 1964. "Measuring Families." Chapter 10 in Handbook of Marriage and the Family, edited by Harold T. Christenson. Chicago: Rand McNally.

_____. 1985a. "The Validity of U.S. States as Units for Sociological Research." Paper presented at the 1985 meeting of the American Sociological Association.

_____. 1985b. "The Index of Legitimate Violence." Unpublished manuscript.

_____. 1987. "Sexual Stratification and Wife Beating in the United States." Paper presented at the Annual Meeting of the American Society of Criminology.

Straus, Murray A. and Bruce W. Brown. 1978. Family Measurement Techniques: Abstracts of Published Instruments, 1935-1974. Minneapolis: University of Minnesota Press.

Straus, Murray A., and Richard J. Gelles. 1980. "Physical Violence in American Families, 1976." Ann Arbor, Michigan: Inter-University Consortium for Political and Social Research, Study No. (ICPSR 7733). (P.O.Box 1248).

_____. 1986. "Societal Change and Change in Family Violence From 1975 to 1985 as Revealed by Two National Surveys." Journal of Marriage and the Family 48 (August): 465-480.

Straus, Murray A., Richard J. Gelles, and Suzanne K. Steinmetz. 1980. Behind Closed Doors: Violence in the American Family. New York: Doubleday/Anchor.

Yllo, Kersti A. 1983a. "Sexual Equality and Violence Against Wives in American States." Journal of Comparative Family Studies XIV:67-86.

_____. 1983b. "Using a Feminist Approach in Quantitative Research: A Case Study." In The Dark Side of Families: Current Family Violence Research, edited by David Finkelhor, Richard J. Gelles, Gerald T. Hotaling, and Murray A. Straus. Beverly Hills, CA: Sage.

_____. 1984. "The Status of Women, Marital Equality, and Violence Against Wives: A Contextual Analysis." Journal of Family Issues 5:307-320.

Yllo, Kersti, and Murray A. Straus. 1984. "Patriarchy and Violence Against Wives: The Impact of Structural and Normative Factors." Paper presented at the Johns Hopkins Symposium on "Feminism and the Critique of Capitalism," Baltimore, MD, April 25, 1981. (An abridged version of this paper was published in The International Journal of Comparative Social Welfare 1 (Fall), 1984.) An unabridged copy of the paper is available from Straus.

Table 1. Concurrent Validity Coefficients (Correlation of Survey Estimates with Census Data) For Five Variables, 1975 National Family Violence Survey

Survey Variable & Census Variable	All 36 States	30 States With N=20+
Median income of husbands by Median income of males	.48	.56
% Husbands Employed full time by % Males employed full time	.13	.24
% Husbands completed High School by % Males 17+ who completed HS	.56	.76
% Wives who completed high school by % Females who completed HS	.68	.77
MEAN VALDITY COEFFICIENT	.46	.58

Table 2. Concurrent Validity Coefficients (r) by Minimum Number of Cases Per State, 1985 National Family Violence Resurvey

	Minimum Number of Cases Per State			
Variable	36 states with N ≥100	39 states with N ≥25	51 states with N ≥7	Mean Correlation
% Age 65	.44	.48	.49B	.47
% Black	.95	.87	.85	.89
% Hispanic	.95	.94	.94	.94
Median Income	.76	.77	.68	.74
Mean Income	.81	.82	.74	.79
MEAN VALDITY COEFF.	.78	.77	.74	.77

Table 3. Conmstruct Valdity Of 1985 National Family Violence Survey variables (correlation with public record state-level variables), for all states and for 36 states with 100+ respondents

NFV Resurvey Variable	Public Record State Level Variable	All 51 States	36 states with N≥100
	A. HOMICIDE RATE		
Approve of H slapping wife under some circumstances (vb49p1)	Homicide rate per 100k pop 1975-80 (vbh4)	.34	.37
Physical Aggression Index of men (vb57hc1)	" "	.13	.40
% black (vbf5p1)	" "	.69	.80
	B. ALCOHOLISM RATE		
% Never drink (vb65ap1)	Alcoholism rate, males 1977 (v112)	-.32	-.57
" "	Alcoholism rate, females, 1977 (v113)2	-.38	-.58
% drinking 3+ days per week (vb65ap2)	Alcoholism rate, males 1977	.20	.49
	Alcoholism rate, females 1977	.28	.53
% husbands drunk 2+ times in last year (vb66ah2)	Alcoholism rate, males 1977	.36	.49
	Alcoholism rate, females 1977	.22	.42
% wives drunk 2+ times in last year (vb67aw2)	Alcoholism rate, males 1977	.19	.57
	Alcoholism rate, females 1977	.32	.58
	C. POVERTY RATE		
Median family income (vbf6m1)	% of children in families with below poverty $ 19?? (cb139rx)	-.27	-.48
% with income below $10,000	" "	.39	.58
	D. STATE STRESS INDEX		
Subjective Stress Index (vb63yp2)	State Stress Index, 1976 (tx15)	.03	.20
Depression Index (vb63xp2)	" "	.29	.30
Psychosomatic Complaints Index (vb63zp1)	" "	.01	.42

Table 4. Concurrent Valdity Coefficients (Correlations Between GSS-Based Variables Nearest Equivalent Census Variable) for Three Sub-sets of States

GSS and Census Variables	All States N=41	States with 50+cases N=34	States with 200+cases N=23
A. Employment			
GS1T4 Percent Laidoff with			
Census CLF Unemployed 1970	.27	.38*	.42*
" " " 1976	.29*	.27	.41*
" " " 1978	.15	.13	.28
GS1T1 Percent Employed Full Time with:			
Census Percent of Pop in the CLF 1970	.56***	.64***	.55**
" " " Employed 1970	.59***	.69***	.58**
" " " in the CLF 1978	.49***	.46**	.48**
GS1T7 Percent Keep House with:			
Census Percent of Female 17 & Work 1975	-.45**	-.41**	-.32
" " " 17 & Work Looked,1975	-.45**	-.43**	-.39*
B. Occupation			
GS2T0 Percent Prof-Tech A with:			
Census Percent of Prof.Manag. 1970	.07	.23	.26
" " " Tech. Kindred 1976	.21	.38*	.38*
GS2T1 Percent Prof-Tech B with:			
Census Percent of Prof. Manag. 1970	.41**	.41**	.53**
" " " Tech. Kindred 1976	.41**	.32*	.52**
GS2T2 Percent Manag.-Admn.-Sales with:			
Census Percent of Prof. Manag. 1970	.28*	.27	.45*
" " Manag. Admn. 1976	.08	.20	.44*
" " Sales 1976	.23	.22	.50**
GS2T3 Percent Clerical with:			
Census Percent Sales-Clerical 1970	.15	.49**	.61***
GS2T4 Percent Craft with:			
Census Percent of Crafts Foremen 1970	.20	.18	.53**
" " " Kindred 1976	.26	.23	.55**
GS2T5 Percent Operatives with:			
Census Percent of Operatives-Trnsprt 1976	-.24	-.32*	-.05
GS2T6 Percent Transport-Labor with:			
Census Percent of Operatives-Trnsprt 1976	.69**	.75***	.66**
" " Laborers non-Farm 1976	.18	.15	.42*
GS2T9 Percent Service with:			
Census Percent of Empl.Service 1976	.19	.26	.49**

Table 4. (Continued) Concurrent Valdity Coefficients for GSS State-Level Variables

C. Marital Status			
GS3T1 Percent Married with:			
Census Percent 14 & Married 1976	.18	.06	.52**
" " " " 1970	.14	.00	.35**
GS3T2 Percent Widowed with:			
Census Percent 14 & Widowed 1976	.32*	.05	.47*
" " " 1970	.37**	.07	.50**
GS3T3 Percent Divorces with:			
Census Percent 14 & Divorces 1976	.53***	.58***	.78***
" " " 1970	.50***	.64***	.81***
GS3T5 Percent Never Married with:			
Census Percent 14 & Never Married 1976	.35*	.37*	.44*
" " Single 1970	.31*	.32*	.31
D. Children			
GS4T0 Percent No Children with:			
Census Percent Families No Children <18 1976	.18	.19	.17
GS4T1 Percent One Child with:			
Census Percent One Child <18 1976	.21	.28	.35
GS4T2 Percent Two Children with:			
Census Percent Two Children <18 1976	.24	.08	.08
GS4T3 Percent Three Children with:			
Census Percent Three Children <18 1976	.42**	.52***	.21
E. Age			
GS5T2 Percent 20-29 Years with:			
Census Percent 20-24 1976	.51***	.63***	.60**
" " 25-34 1976	.39**	.50**	.64***
GS5T3 Percent 30-39 Years with:			
Census Percent 25-34 1976	.09	-.02	.18
" " 35-44 1976	.14	-.02	.18
GS5T4 Percent 40-49 Years with:			
Census Percent 35-44 1976	-.06	.11	-.16
" " 45-54 1976	.15	.22	.20
GS5T5 Percent 50-59 Years with:			
Census Percent 45-54 1976	.22	.23	.32
" " 55-64 1976	.10	.23	.31
GS5T6 Percent 60-69 Years with:			
Census Percent 55-64 1976	.33*	.33*	.29
" " 65+ 1976	.31*	.24	.35
GS5T7 Percent 70-79 Years with:			
Census Percent 65+ 1976	.31*	.32*	.37*

Table 4. (Continued) Concurrent Valdity Coefficients for GSS State-Level Variables

GS5T8 Percent 80+ Years with:			
Census Percent 65+ 1976	.22	.29*	.47*
F. Education			
GS6T1 Percent 0-7 Years Education with:			
Census Percent 0-4 Years Education 1976	.75***	.81***	.79***
" " 5-8 " " "	.46**	.57***	.50**
GS6T2 Percent 8 Years Education with:			
Census Percent 5-8 Years Education 1976	.56***	.44**	.49**
GS6T3 Percent 9-11 Years Education with:			
Census Percent 1-3 Years H.S. 1976	.51***	.56***	.64***
GS6T4 Percent 12 Years Education with:			
Census Percent 4 Years H.S. 1976	.51***	.74***	.91***
GS6T5 Percent 1-3 Years College with:			
Census Percent 1-3 Years College 1976	.58***	.72***	.71***
" " " " " 1970	.55***	.67***	.66***
GS6T6 Percent 4 Year College with:			
Census Percent 4+ Years College 1976	.46**	.46**	.48**
GS6T7 Percent 5-8 Years College with:			
Census Percent 4+ Years College 1970	.33*	.44**	.57**
G. Gender and Race			
GS7T2 Percent Female with:			
Census Males per 100 Females 1970	-.08	.08	-.14
GS8T2 Percent Black with:			
Census Percent Black 18+ 1976	.64***	.65***	.77***
" " " All Ages 1976	.66***	.67***	.76***
H. Household Composition			
GS9T1 Percent in 1-Person H.H.'s with:			
Census Percent H.H. 1-Person 1976	.22	.31*	.43*
GS9T2 Percent in 2-Person H.H.'s with:			
Census Percent Families 2-Member 1976	.18	.22	.50**
GS9T3 Percent in 3-Person H.H.'s with:			
Census Percent Families 3-Member 1976	.04	.40**	.28
MEAN CORRELATION COEFFICIENT	.34	.36	.45

* - p <.05, ** - p <.01, *** - p <.001 CLF - Civilian Labor Force